

Fig. S1

A	<u>MKTEEGKLVI</u>	<u>WINGDKGYNG</u>	<u>LAEVGKKFEK</u>	<u>DTGIKVTVEH</u>	<u>PDKLEEKFPQ</u>	<u>VAATGDGPPDI</u>	60
	<u>IFWAHDRFGG</u>	<u>YAQSGLLAEI</u>	<u>TPDKAFQDKL</u>	<u>YPFTWDVARY</u>	<u>NGKLIAYPIA</u>	<u>VEALSIIYNK</u>	120
	<u>DLLPNPPKWTW</u>	<u>EEIPALDKEL</u>	<u>KAKGKSALMF</u>	<u>NLQEPYFTWP</u>	<u>LIAADGGYAF</u>	<u>KYENGKYDIK</u>	180
	<u>DVGVDNAGAK</u>	<u>AGLTFLVDLI</u>	<u>KNKHMNADTD</u>	<u>YSIAEAAFNK</u>	<u>GETAMTINGP</u>	<u>WAWSNIDTSK</u>	240
	<u>VNYGVTVLPT</u>	<u>FKGQPSKPFV</u>	<u>GVLSAGINAA</u>	<u>SPNKELAKEF</u>	<u>LENYLLTDEG</u>	<u>LEAVNKDKPL</u>	300
	<u>GAVALKSYEE</u>	<u>ELAKDPRIAA</u>	<u>TMENAQKGEI</u>	<u>MPNIPQMSAF</u>	<u>WYAVRTAVIN</u>	<u>AASGRQTVDE</u>	360
	<u>ALKDAQTNSS</u>	<u>SNNNNNNNN</u>	<u>NLGIEGRISE</u>	<u>FGSTSRVDCG</u>	<u>GLTGLNSGLT</u>	<u>TNPGVSAWQV</u>	420
	<u>NTAYTAGQLV</u>	<u>TYNGKTYKCL</u>	<u>QPHTSLAGWE</u>	<u>PSNVFALWQL</u>	<u>Q</u>		461

B	[M+H] ⁺		[M+H] ⁺		Mass	
	Peptide position	calculated mass	measured mass	ppm	accuracy	Sequence
	1-2	278.15	278.14	36		MK
	3-7	563.27	563.26	18		TEEGK
	8-16	1057.60	1057.60	0		LVIWINGDK
	8-26	2046.10	2046.23	64		LVIWINGDKGYNGLAEVGK
	27-35	1065.59	1065.59	0		KFEKDTGIK
	28-30	423.22	423.21	24		FEK
	36-47	1423.74	1422.73	709		VTVEHPDKLEEK
	90-99	1267.65	1267.60	39		LYPFTWDAVR
	129-138	1201.61	1201.60	8		TWEEIPALDK
	129-141	1571.83	1571.82	6		TWEEIPALDKELK
	191-201	1189.72	1189.71	8		AGLTFLVDLIK
	172-180	1129.55	1129.57	18		YENGKYDIK
	253-274	2139.16	2139.08	37		GQPSKPFVGVLSAGINAAASPDK
	279-296	2097.04	2096.99	24		EFLENYLLTDEGLEAVNK
	297-306	1011.62	1011.57	49		DKPLGAVALK
	307-317	1336.64	1336.63	7		SYEEELAKDPR
	328-345	2110.03	2110.01	9		GEIMPNI PQMSAFWYAVR
	346-355	959.53	959.52	10		TAVINAAAGR
	356-387	3459.57	3458.61	277		QTVDEALKDAQTNSSNNNNNNNNNNNLGIEGR
	364-387	2575.11	2575.41	116		DAQTNSSNNNNNNNNNNNLGIEGR
	388-396	983.48	983.55	71		<u>ISEFGSTSR</u>
	397-435	3985.92 ^a	3985.88	10		VDCGGLTGLNSGLTTNPGVSAWQVNTAYTAGQLVTYNG
	397-438	4378.13 ^a	4378.21	18		VDCGGLTGLNSGLTTNPGVSAWQVNTAYTAGQLVTYNGKTYK
	439-461	2633.28 ^a	2633.14	53		CLQPHTSLAGWEPSNVPALWQLQ

Fig. S1. Assignment of MALDI Peptide Mass Map to the MC ligation product (Fig. 6A).

A. Sequences detected by MALDI analysis of the MC product are underlined. Twenty-five tryptic peptide masses were assigned to the amino acid sequence of the MC protein, corresponding to sequence coverage of 49%. Amino acids matching the C-tag protein are in italic. The double underlined peptide (ISEFGSTSR - amino acids 388–396) contains the BIL splice site between amino acids Ser393 and Thr394.

B. Measured and calculated masses for tryptic peptides which identify the 50.6 kD MC protein.

^a mass corresponds to peptide with an alkylated cysteine

Fig. S2

A

<u>MKTEEGKLVI</u>	<u>WINGDKGYNG</u>	<u>LAEVGKKFEK</u>	<u>DTGIKVTVEH</u>	<u>PKLEEKFPQ</u>	<u>VAATGDGPDI</u>	60
<u>IFWAHDRFGG</u>	<u>YAQSGLLAEI</u>	<u>TPDKAFQDKL</u>	<u>YPFTWDAVRY</u>	<u>NGKLIAYPIA</u>	<u>VEALSIIYNK</u>	120
<u>DLIPNPPTW</u>	<u>EEIPALDKEL</u>	<u>KAKGKSALMF</u>	<u>NLQEPYFTWP</u>	<u>LIAADGGYAF</u>	<u>KYENGKYDIK</u>	180
<u>DVGVDNAGAK</u>	<u>AGLTFLVDLI</u>	<u>KNKHMNADTD</u>	<u>YSIAEAAFNK</u>	<u>GETAMTINGP</u>	<u>WAWSNIDTSK</u>	240
<u>VNYGVTVLPT</u>	<u>FKGQPSKPFV</u>	<u>GVLSAGINAA</u>	<u>SPNKELAKEF</u>	<u>LENYLLTDEG</u>	<u>LEAVNKDKPL</u>	300
<u>GAVALKSYEE</u>	<u>ELAKDPRIAA</u>	<u>TMENAQKGEI</u>	<u>MPNIPQMSAF</u>	<u>WYAVRTAVIN</u>	<u>AASGRQTVDE</u>	360
<u>ALKDAQTNSS</u>	<u>SNNNNNNNNN</u>	<u>NLGIEGRISE</u>	<u>FGscfaagtm</u>	<u>vstpdgerai</u>	<u>dtlkvqdivw</u>	420
<u>skpegggkpf</u>	<u>aaailathir</u>	<u>tdqpiyrkl</u>	<u>kgkqenggae</u>	<u>desllvtppgh</u>	<u>pfyvpaghgf</u>	480
<u>vpvidlkgpd</u>	<u>rlqsladgas</u>	<u>entssevesl</u>	<u>elylpvgkty</u>	<u>nltvdvghtf</u>	<u>yvgklktwvh</u>	540
<u>n</u>						541

B

Peptide position	[M+H] ⁺ calculated mass	[M+H] ⁺ measured mass	Mass accuracy ppm	Sequence
535-541	897.4947	897.47	27.5	LKTWVHN
537-541	656.3156	56.25	99.9	TWVHN

Fig. S2. MALDI peptide mapping of the 59.3 kD MB protein (Fig. 6A).

A. Underlined sequences correspond to peptides detected by MALDI. Uppercase letters match amino acids of the M-tag and lowercase letters match those of the BIL domain. Note that the C-terminus of the protein, Asn 541, is the penultimate C-terminal residue of the BIL sequence (Fig. 4).

B. Measured and calculated molecular masses of the two C-terminal peptides.